Vol.6, No.2, October 2020 Available online at: http://e-journal.unair.ac.id/index.php/JISEBI

Tweets Responding to the Indonesian Government's Handling of COVID-19: Sentiment Analysis Using SVM with Normalized Poly Kernel

Pulung Hendro Prastyo^{1)*}, Amin Siddiq Sumi²⁾, Ade Widyatama Dian³⁾, Adhistya Erna Permanasari⁴⁾

¹⁾²⁾³⁾⁴⁾Department of Electrical and Information Technology Engineering, Faculty of Engineering, Universitas Gadjah Mada, Indonesia

Jalan Grafika No.2 Kampus UGM, Sleman, Yogyakarta

¹⁾pulung.hendro@mail.ugm.ac.id, ²⁾aminsiddiqsumi@mail.ugm.ac.id, ³⁾adewidyatama@mail.ugm.ac.id, ⁴⁾adhistya@ugm.ac.id

Article history:	Abstract
Received 4 June 2020 Revised 9 September 2020 Accepted 10 September 2020 Available online 28 October 2020	Background: Handling COVID-19 (Corona Virus Disease-2019) in Indonesia was once trending on Twitter. The Indonesian government's handling evoked pros and cons in the community. Public opinions on Twitter can be used as a decision support system in making appropriate policies to evaluate government performance. A sentiment analysis method can
Keywords:	be used to analyse public opinion on Twitter. Objective: This study aims to understand public opinion trends on COVID-19 in Indonesia
COVID-19 Economic Aspect Predictive Analysis Sentiment Analysis Support Vector Machine Text Mining	both from a general perspective and an economic perspective. Methods: We used tweets from Twitterscraper library. Because they did not have a label, we provided labels using sentistrength_id and experts to be classified into positive, negative, and neutral sentiments. Then, we carried out a pre-processing to eliminate duplicate and irrelevant data. Next, we employed machine learning to predict the sentiments for new data. After that, the machine learning algorithms were evaluated using confusion matrix and K-fold cross-validation.
	 Results: The SVM analysis on the sentiments on general aspects using two-classes dataset achieved the highest performance in average accuracy, precision, recall, and f-measure with the value of 82.00%, 82.24%, 82.01%, and 81.84%, respectively. Conclusion: From the economic perspective, people seemed to agree with the government's policies in dealing with COVID-19; but people were not satisfied with the government performance in general. The SVM algorithm with the Normalized Poly Kernel can be used as an intelligent algorithm to predict sentiment on Twitter for new data quickly and accurately.

I. INTRODUCTION

The use of the Internet, especially the social media, is high in Indonesia. Based on the results of Hootsuite Social Wear research released in January 2019, social media users in Indonesia reached 150 million or 56% of the total population, increasing by 20% from the previous survey[1]. One of the widely-used and growing social media is Twitter[2]. Content shared by Twitter users is mostly opinions on current issues such as the government policies and the socio-political issues; as well as new products' reviews (which can be used for business intelligence). Either way, the content can be used as an effective source of information to help a decision-making process.

The COVID-19 pandemic has impacted all sectors and communities in Indonesia. According to data from covid19.go.id, on 14 May 2020, there were 16.006 confirmed cases, 3.518 patients recovered, and 1.043 patients died [3]. The Indonesian government is considered slow in handling COVID-19 [4][5] and so the Indonesian government's policy in dealing with the COVID-19 outbreak became a trending topic on Twitter, evoking the pros and cons, especially because COVID-19 has been impacting many sectors severely.

The government policy must be well informed to be on target. Public opinions such as those published on Twitter are needed to inform the policy making. But extracting public opinions on Twitter is not easy, because they are messy and are often written in non-standard languages. If these are processed adequately, knowledge can be

ISSN 2443-2555 (online) 2598-6333 (print) © 2020 The Authors. Published by Universitas Airlangga. This is an open access article under the CC BY license (<u>http://creativecommons.org/licenses/by/4.0/</u>) doi: http://dx.doi.org/10.20473/jisebi.6.2.112-122

^{*} Corresponding author

obtained to assist policy making. Therefore, we need a particular method or technique that can classify the opinions quickly.

Sentiment analysis in text mining studies sentiments, emotions, and attitudes in opinion texts. The basic principle of sentiment analysis is to classify the polarity of a given text and determine whether the content is more positive, neutral or negative[6]. Therefore, this study uses sentiment analysis to examine the responses towards the government's handling of COVID-19. The data in this study are divided into general and economic related sentiments. The general perspective aims to understand Indonesian opinions in a global sense, such as health, education, facilities, etc.; while the economic perspective looks into more details how the government solutions to the pandemic through its fiscal and economic policies were responded by the public [7].

Sentiment analysis research has been developing rapidly and there have been new approaches launched. Machine Learning-based sentiment analysis has been a popular approach considered to be effective in analysing text-based data[8]. In previous studies, Kristiyanti et al. [9] conducted a sentiment analysis of the public opinions towards the West Java governor candidate in the period of 2018-2023. The researchers compared two algorithms: the Support Vector Machine (SVM) and Naïve Bayes Classifier (NBC). The results showed that NBC had a higher accuracy than SVM.

G. Singh et al.[10] compared Multinomial Naïve Bayes (MNB) and Bernoulli Naïve Bayes (BNB) to predict whether a particular news article's sentiment is positive or negative. The experimental results showed that MNB performed better than BNB. A. Lutfi et al.[11] conducted a sentiment analysis in the sales review of the Indonesian market place by employing SVM and NBC; and the results show that SVM outperformed NBC as it obtained 93.65% of accuracy. Other researchers supported the results stating that SVM and MNB had a good performance to classify opinions [12]–[19]

In SVM, choosing kernel functions is essential for an effective SVM based classification because it will provide the necessary learning capabilities. The use of applicable kernel functions can drastically reduce computational efforts to make operations feasible. Researchers have employed kernel of SVM in their studies, such as Polynomial Kernel (PK) [20]–[22], Radial Basis Function Kernel (RBF) [20]–[23], Gaussian Kernel [21], Linear Kernel [21], Sigmoid Kernel [21], Laplacian Kernel [21], and Anova Kernel [21], but there are only a few researchers applying Normalized Poly Kernel despite its excellent performance [24][25]. Thus, in this research, we employed a Normalized Poly Kernel in Support Vector Machine (SVM)

Finally, in this study, we develop machine learning algorithms and compared SVM using Normalized Poly Kernel and MNB to determine which algorithm works best in conducting sentiment analysis of the Indonesian government's handling of COVID-19. We also show the word cloud to find out what words appear most often in this problem. This study is organized as follows: Section II presents the related work of sentiment analysis; Section III explains the methods; Section IV describes results and research discussions; and Section V concludes this study with a summary and direction for future work.

II. RELATED WORKS

Past research using machine learning to conduct sentiment analysis utilized different algorithms. M. Pratama et al. [26], analyzed the sentiments towards the Indonesia Commuter Line (KRL) using Twitter data. They classified public opinions into positive and negative classes. Pre-processing was used to remove irrelevant data that did not represent any sentiment such as symbols, numbers, punctuations and links. The data processing involved stemming, normalization and removing stop words as well. Multinomial Naive Bayes (MNB), Random Forest (RF) and Support Vector Machine (SVM) were used as a model. The results showed that SVM was superior to the other algorithms and obtained 85% accuracy.

In politics, Krisyanti et al.[9] conducted sentiment analysis towards the West Java governor candidate in 2018-2023 using SVM and Naïve Bayes (NB). They examined 800 tweets as datasets and used tokenization and generated n-gram in the pre-processing. SVM and NB were compared to determine which model performed better. Ten-folds cross-validation was used to validate the models. Confusion matrix and ROC curve were employed to measure the models. The experimental results showed that NB obtained 94% of accuracy and SVM 75.50 %.

G. Singh [10] examined the sentiment analysis in news articles. They pre-processed 312 articles using lowercasing, tokenization, punctuation removal, and stop word removal. Multinomial Naïve Bayes (MNB) and Bernoulli Naïve Bayes (BNB) were used as the classifier. The results showed that MNB achieved 73% accuracy and slightly better than BNB.

A. Lutfi et al. [11] examined datasets from a marketplace Bukalapak, one of the widely-used online stores in Indonesia. They used 3177 reviews, consisting of 1521 negative reviews and 1656 positive reviews. The data were labeled as positive or negative class manually. They divided the data into training data and testing data. Before

employing the algorithms, they pre-processed the data using case folding, stemming, tokenization, slang words identification, negation, and stop words as well as feature extraction (TF-IDF). SVM and NB were employed to classify reviews. Ten-folds cross-validation was used to validate the built model, and the accuracy was measured to determine the number of the documents to be classified correctly by the system. The results showed that SVM with a linear kernel was superior to NB as it provided 93.65 % average accuracy using 25% features with the highest TF-IDF.

M. Zul et al. [18] presented sentiment analysis on social media using K-Means and Naïve Bayes algorithm. They compared Naïve Bayes and a combination of Naïve Bayes and K-Means. Data were collected from Facebook and Twitter using Power Query and Twitter API. In the pre-processing, the data were treated using case folding, negation word, deletion of punctuation, tokenizing, stop word removal, and stemming. Outlier detection was accomplished using Knime. The results showed that Naïve Bayes obtained better accuracy than the combination of Naïve Bayes and K-Means.

L. Muflikhah et al. [20] compared the performance of Polynomial and Radial Basis Function (RBF) kernel in SVM in the sentiment analysis of online product reviews. Data were collected from 200 comments and preprocessed using tokenization, stop word removal, stemming, and normalization in data preprocessing. They also employed various parameter values i.e. learning rate, lambda, C value, epsilon, and iteration and validated the data using ten-fold validation. The results showed that the performance of Polynomial kernel outperformed RBF.

Trivedi and Dey [24] examined the effect of various kernels and feature selection methods on SVM performance in detecting email spams. Four kernel functions of SVM were compared, i.e. Normalized Polynomial Kernel (NP), Polynomial Kernel (PK), Radial Basis Function Kernel (RBF), and Pearson VII Function-based Universal Kernel (PUK). The results showed that NP outperformed the other kernels, obtaining 98.5% accuracy for the 1358 high dimensional data features and 78.1% to 85.2% accuracy for the 158 low dimensional data features.

III. METHODS

In the current research, Indonesian tweets were classified into two types of data. The first dataset consists of three classes (positive, neutral, and negative class) and the second dataset consists of two classes (positive and negative class). Several stages were required to obtain an optimum classification result, namely tweets scraping, labeling, preprocessing, feature extraction, development of the classification model, evaluation and word cloud. The framework can be seen in Fig 1.



Fig. 1 The research framework

A. Data Collection

Data collection was started by scraping on the site twitter.com from 23 March 2020 to 14 May 2020 using the Twitterscraper library [27]. This library can retrieve data without a time limit, unlike Twitter API that can only retrieve data for the past week. Twitterscrapper library uses keywords in the form of sentence, word, or hashtag to retrieve data from Twitter. The results of scraping data using Twitterscrapper library were 20,444 tweets for general aspects using keywords "#COVID-19indonesia, COVID-19 di Indonesia, penanganan pemerintah terhadap COVID-19" and 14,451 tweets for economic aspects using keywords "Dampak ekonomi karena COVID-19 di Indonesia, penanganan ekonomi COVID-19 di Indonesia". Upon the data collection, irrelevant and duplicated data were removed so that the total data became 2.203 tweets for the general aspects and 1.941 tweets for the economic aspects. Then, the data were labeled using sentistrength_id [28] and validated by the experts. The labeling used two datasets. The first dataset consists of three classes (positive, neutral, and negative) and the second dataset consists of two classes and three classes. This was carried out in the experiments section and presented on the experimental results.

TABLE 1	
EXAMPLE OF PRE-PROCESSING STAGES AND THE RESUL	1

Stage	Tweet	Class
Initial Tweet	@RamliRizal Saat ini Pemerintah, TNI @Puspen_TNI & POLRI @DivHumas_Polri, Media @PBIDI sdng sangat Focus memutuskan rantai penularan CoVID-19;	Positive
Casefolding	@ramlirizal saat ini pemerintah, tni @puspen_tni & polri @divhumas_polri, media @pbidi sdng sangat focus memutuskan rantai penularan covid-19;	Positive
Filtering	saat ini pemerintah tni polri media sdng sangat focus memutuskan rantai penularan covid-19	Positive
Tokenization	[saat] [ini] [pemerintah] [tni] [polri] [media] [sdng] [sangat] [focus] [memutuskan] [rantai] [penularan] [covid-19]	Positive
Stopword Removal	pemerintah tni polri media focus memutuskan rantai penularan covid-19	Positive
Clean tweet	pemerintah tni polri media focus memutuskan rantai penularan covid-19	Positive

B. Pre-processing Data

The textual data from Twitter are messy so we ran them through pre-processing stages to produce relevant tweet data and to ease the performance of the text classification algorithms. The pre-processing stages are as follows. The results are shown in Table 1.

1) Case folding

At this stage, the tweets were converted into all lowercase by using Python programming language, with the function of lower().

2) Filtering

We removed strings like @username, website (https:// or http:// or www), picture (pic.twitter.com); and symbols such as punctuation e.g. commas (,), periods (.), question marks (?), exclamation points (!) and so on; numeric (0 - 9); and other characters e.g. %, * and so on .

3) Tokenization

Tokenization is the process of dividing tweets into words by using the NLTK library in Python; for example, "*wabah virus corona tidak akan berakhir jika pemerintah tidak melarang warga china tetap masuk ke Indonesia*" to "[wabah] [virus] [corona] [tidak] [akan] [berakhir] [jika] [pemerintah] [tidak] [melarang] [warga] [china] [tetap] [masuk] [ke] [Indonesia]".

4) Stopword Removal

At this stage, we took essential words from the tokenization results by using a stop list algorithm (removing less important words) or wordlist (save essential words), carried out by the Sastrawi library [29] using the stopWordRemoverFactory module.

C. Feature Extraction

Term weighting TF-IDF (Term Frequency-Inverse Document Frequency) was employed to extract and convert words into vectors. TF measures how frequently a term appears in a document. Meanwhile, IDF computes how important a term is. It does not consider words that appear a lot of times but have little importance, such as *ini*, *di*, *ke*, *dan*, *sana*, soon [30].

TF-IDF is a weighting method that combines the two methods, which can improve the performance, especially in improving the value of recall and precision [31]. The TF-IDF method calculates each word's value in a document using the frequency of occurrence of words [31]. The TF equation can be seen in Equation (1).

$$TF_{(d,t)} = \frac{f_{t,d}}{\sum t,d} \tag{1}$$

 $f_{t,d}$ = The frequency of each word (t) appearing in document d and $\sum t, d$ = Total of all words contained in document d. for IDF value, it can be calculated using equation (2) as follows:

$$IDF_{(t)} = \log \frac{|D|}{f_{t,d}} \tag{2}$$

Where, |D| = Number of documents in the collection and $f_{t,d} =$ Number of documents where (t) appears in document d. Then the TF-IDF (W) value can be calculated using equation (3) as follows:

$$W_{(d,t)} = TF_{(d,t)} \cdot IDF_{(t)}$$
(3)

D. Development of Classification Model

After the pre-processing and the feature extraction, we developed machine learning algorithms that were subsequently used to classify unseen data. We used 2,203 tweets for general aspects: 787 positive, 482 neutral, and 934 negative sentiments. As for the economic aspects, we used 1,941 tweets consisting of 973 positive, 385 neutral, and 585 negative sentiments. We then trained the data using machine learning algorithms, namely Multinomial Naïve Bayes (MNB) and Support Vector Machine (SVM) with Normalized Poly Kernel. In this experiment, we used two and three classes to assess the algorithm in the different classes.

SVM is a supervised-learning algorithm frequently used in classification and regression. It works by looking for a hyper-plane, whose best location is in the center between two classes. SVM finds the best hyper-plane equation to maximize the distance between two groups in different classes [26]. An essential thing for SVM is the choice of the kernel function. To obtain the learning ability, SVM needs a suitable kernel function. Therefore, we use SVM with Normalized Poly kernel (NP) using Weka because NP performs better than other kernels [24][25].

Meanwhile, Multinomial Naïve Bayes (MNB) is a modified Naïve Bayes classifier aiming to achieve better performance than the prior method. MNB misclassifies values less frequently than the Naïve Bayes algorithm [30]. Without having to recognize the sentence multiple times in a document, MNB can still process information correctly. Besides, it only requires a small amount of data training to evaluate the classification parameters, and it is robust to noise in input data. In brief, MNB is an effective algorithm to overcome classification problems [32]. Thus, we used both classifiers and evaluated their performance. We used Weka and Jupyter in Python to perform both of algorithms.

E. Evaluation and Validation

We used a confusion matrix in the evaluation process. The confusion matrix is a tool to evaluate the performance of machine learning algorithms that contain information about the actual classification and prediction. There are four indicators measured in the confusion matrix: accuracy, recall, precision, and F1 score [33]. The scenario can be seen in Fig. 2.

		Predicted class				
		Class = True	Class = False			
Actual class	Class = True	True positive	False Negative			
Class	Class = False	False Positive	True Negative			

Fig. 2 Confusion Matrix

The accuracy was calculated using Equation (5).

The recall was calculated using Equation (6).

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)}$$
(5)

$$Recall = \frac{TP}{(TP+FN)} \tag{6}$$

The precision was calculated using Equation (7).

 $Precision = \frac{TP}{(TP+FP)}$ (7)

The F1 score was calculated using Equation (8).

$$F1 - Measure = 2 \times \frac{precision \times recall}{precision + recall}$$
(8)

After that, the algorithms were validated using K-folds cross-validation to understand the variation of results against the built model to avoid overfitting. K-folds cross-validation divides training and testing data iteratively as many as K values to test the entire data [33][34]. The K-fold cross-validation scheme can be seen in Fig 3.



Fig. 3 K-folds cross-validation

IV. RESULTS

As mentioned previously, the tweet data was divided into two datasets: the first set consists of two classes (positive and negative) and the second consists of three classes (positive, neutral, and negative). This applied to both on general sentiments and the economic sentiments. The classification aimed to determine the ability of the algorithms to classify and predict new data or unseen data. SVM with Normalized Poly Kernel, and Multinomial Naïve Bayes (MNB) were used as classifiers. The results were compared to determine which algorithm worked best.

A. Result of Preprocessing and Labeling

Based on the research results, we obtained 1.941 cleaned tweets for the economic aspects consisting of 973 positive, 385 neutral, and 585 negative sentiments. The results showed that the public opinions on the economic policies in Indonesia tended to be positive. The graph of the sentiments related to the economic aspects can be seen in Fig. 4.



Fig. 4 The graph of sentiment in the economic aspect

As for general aspects, from the cleaned 2.203 tweets, the results showed that of 787 were positive, 482 neutral, and 934 negative. In general, the public seemed to be unsatisfied with the government's handling of COVID-19. The general aspects are health, aid distribution, government's response, general policies, and others. The graph of sentiments related to general aspects can be seen in Fig. 5.



Fig. 5 The graph of sentiment in the general aspect

CLASSIFICATION RESULTS OF THE ECONOMIC ASPECT USING THREE CLASSES									
		М	NB		SVM				
10-Folds	Three C	lasses (Positi	ive, Neutra	il, Negative)	Three C	lasses (Positi	ive, Neutra	l, Negative)	
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	
1	56,11	58,10	56,10	56,90	64,92	63,30	64,90	61,50	
2	56,52	58,30	56,50	57,20	65,89	64,60	65,90	62,70	
3	56,05	58,40	56,10	57,00	65,84	64,60	65,80	62,40	
4	56,77	59,00	56,80	57,60	66,20	64,80	66,20	62,90	
5	56,93	58,90	56,90	57,70	65,95	64,60	65,90	62,70	
6	56,05	57,80	56,10	56,80	66,82	65,40	66,80	63,30	
7	56,11	58,40	56,10	57,00	65,64	64,70	65,60	62,40	
8	57,34	59,80	57,30	58,30	65,89	63,90	65,90	62,50	
9	57,86	60,30	57,90	58,80	65,69	64,60	65,70	62,40	
10	57,24	59,20	57,20	58,00	65,69	64,30	65,70	62,30	
Minimum	56,05	57,80	56,10	56,80	64,92	63,30	64,90	61,50	
Maximum	57,86	60,30	57,90	58,80	66,82	65,40	66,80	63,30	
Mean	56,70	58,82	56,70	57,53	65,85	64,48	65,84	62,51	
Standard Deviation	0,64	0,78	0,63	0,67	0,48	0,56	0,48	0,47	

 TABLE 2

 CLASSIFICATION RESULTS OF THE ECONOMIC ASPECT USING THREE CLAS

B. Result of Classification Model Based on Sentiment in The Economic Aspect

Based on the experimental results using three classes (positive, neutral, and negative), the SVM algorithm outperformed MNB with an average accuracy of 65.85%, precision of 64.48%, recall of 65.84%, and f-measure of 62.51%. The value differences of each measurement are by 9.15%, 5.66%, 9.14%, and 4.98%, respectively. Detailed experimental results can be seen in Table 2.

For the experimental results using two classes (positive and negative), the SVM algorithm still outperformed MNB with an average accuracy of 80.08%, the precision of 79.91%, the recall of 80.08%, and the f-measure of 79.65%. Interestingly, using two classes (positive and negative), both SVM's performance and MNB's performance rose. SVM increased accuracy, precision, recall, and f-measure by 14.23%, 15.43%, 14.68%, and 17.14%, respectively. SVM achieved a maximum accuracy of 81.04%. Besides, MNB also increased significantly. This means that both algorithms perform better in predicting two classes than in predicting three classes. Detailed experimental results using two classes of sentiments in the economic aspects can be seen in Table 3.

10-Folds	Т	N wo Classes (I	INB Positive, N	egative)	Т	SVM Two Classes (Positive, Negative)			
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	
1	75,00	75,20	75,00	75,10	79,50	79,30	79,50	79,00	
2	74,16	74,40	74,20	74,30	79,88	79,70	79,90	79,50	
3	74,61	75,00	74,60	74,80	80,27	80,20	80,30	79,80	
4	74,29	74,70	74,30	74,50	80,33	80,20	80,30	79,80	
5	74,94	75,30	74,90	75,10	80,08	79,90	80,10	79,60	
6	74,04	74,44	74,00	74,20	79,63	79,40	79,60	79,20	
7	74,94	75,30	74,90	75,10	81,04	80,90	81,00	80,70	
8	75,45	75,80	75,40	75,60	80,21	80,00	80,20	79,80	
9	74,94	75,30	74,90	75,10	80,53	80,40	80,50	80,10	
10	74,55	74,90	74,60	74,70	79,37	79,10	79,40	79,00	
Minimum	74,04	74,40	74,00	74,20	79,37	79,10	79,40	79,00	
Maximum	75,45	75,80	75,40	75,60	81,04	80,90	81,00	80,70	
Mean	74,69	75,03	74,68	74,85	80,08	79,91	80,08	79,65	
Standard Deviation	0,44	0,44	0,42	0,43	79,50	79,30	79,50	79,00	

 TABLE 3

 CLASSIFICATION RESULTS OF THE ECONOMIC ASPECT USING TWO CLASSES

TABLE 4

CLASSIFICATION RESULTS OF THE GENERAL ASPECT USING THREE CLASSES									
		M	NB			SV	/M		
10-Folds	Three C	lasses (Positi	ve, Neutra	l, Negative)	Three C	Three Classes (Positive, Neutral, Negative)			
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure	
1	62,41	63,10	62,40	62,70	67,77	68,40	67,80	65,70	
2	61,96	62,50	62,00	62,20	67,59	68,00	67,60	65,50	
3	62,87	63,70	62,90	63,20	69,22	70,20	69,20	67,30	
4	62,14	63,10	62,10	62,50	68,63	69,20	68,60	66,50	
5	62,69	63,40	62,70	63,00	68,18	69,00	68,20	66,20	
6	61,73	62,30	61,70	62,00	68,27	68,80	68,30	66,30	
7	62,91	63,50	62,90	63,20	68,00	68,60	68,00	66,00	
8	63,28	63,39	63,30	63,50	67,82	68,40	67,80	65,80	
9	61,68	62,50	61,70	62,00	68,00	68,30	68,00	65,70	
10	62,01	62,70	62,00	62,30	68,13	68,70	68,10	65,90	
Minimum	61,68	62,30	61,70	62,00	67,59	68,00	67,60	65,50	
Maximum	63,28	63,70	63,30	63,50	69,22	70,20	69,20	67,30	
Mean	62,37	63,02	62,37	62,66	68,16	68,76	68,16	66,09	
Standard Deviation	0,55	0,49	0,56	0,54	0,47	0,61	0,46	0,52	

 TABLE 5

 CLASSIFICATION RESULTS OF THE GENERAL ASPECT USING TWO CLASSES

		N	INB		SVM			
10-Folds	Tw	vo Classes (P	ositive, Ne	gative)	Two Classes (Positive, Negative)			
	Accuracy	Precision	Recall	F-Measure	Accuracy	Precision	Recall	F-Measure
1	80,36	80,30	80,40	80,30	81,81	82,10	81,80	81,60
2	79,31	79,30	79,30	79,30	81,87	82,10	81,90	81,70
3	79,08	79,10	79,10	79,10	81,81	82,10	81,80	81,60
4	79,02	79,00	79,00	79,00	82,28	82,50	82,30	82,10
5	79,02	79,00	79,00	79,00	81,58	81,80	81,60	81,40
6	78,91	78,90	78,90	78,90	82,28	82,40	82,30	82,20
7	78,97	79,00	79,00	79,00	82,57	82,80	82,60	82,40
8	79,66	79,60	79,70	79,70	82,74	83,00	82,70	82,60
9	78,73	78,70	78,70	78,70	82,16	82,40	82,20	82,00
10	78,27	78,20	78,30	78,30	80,94	81,20	80,90	80,80
Minimum	78,27	78,20	78,30	78,30	80,94	81,20	80,90	80,80
Maximum	80,36	80,30	80,40	80,30	82,74	83,00	82,70	82,60
Mean	79,13	79,11	79,14	79,13	82,00	82,24	82,01	81,84
Standard Deviation	0,56	0,55	0,57	0,55	0,52	0,51	0,53	0,53

C. Result of Classification Model Based on Sentiment in The General Aspect

The sentiment in the general aspects using three classes (positive, neutral, and negative) showed that the SVM algorithm was better than MNB. SVM obtained the average accuracy of 68.16%, the precision of 68.76%, the recall of 68.16%, and the f-measure of 66.09%. The difference in the SVM performance against MNB in accuracy, precision, recall, and f-measure was by 5.79%, 5.74%, 5.79%, and 3.43%, respectively. Detailed experimental results can be seen in Table 4.

For the experimental results using two classes (positive and negative), the performance of the two algorithms significantly improved. The SVM algorithm was superior to MNB. The SVM achieved the best performance in the average accuracy, precision, recall, and f-measure with the value of 82.00%, 82.24%, 82.01%, and 81.84%, respectively. Both algorithms can classify tweet data better in two-classes dataset rather than three-classes dataset. Detailed experimental results using two classes on the general-aspect sentiments can be seen in Table 5.

D. Sentiment of The Economic and General Aspects in World Cloud

The word cloud shows the most frequently appeared words in the data, which are pictured bigger than the other words. On the economic aspects, the words that appeared most often were *dampak ekonomi*, *dampak covid*, *virus corona*, *pemerintah*, *Indonesia*, *sosial ekonomi*, and *dampak sosial*. Positive sentiments on economic aspects were more significant than the negative sentiments. It means that the public opinions on Twitter tend to agree and support government economic policies in the handling COVID-19.

As for the general aspects (health, facilities, government performance, etc.), the sentiments tend to be negative by showing the words that appeared most frequently, namely *pemerintah*, *lockdown*, *Indonesia*, *rakyat*, *masyarakat*, *covid*, and *covid Indonesia*. It means that the public opinions tend to be negative with government general policies in dealing with COVID-19. The sentiments of the economic and general aspects in word cloud can be seen in Fig. 6.



(a) Sentiment of the economic aspect



(b) Sentiment of the general aspect

Fig. 6 The word cloud of sentiment

V. DISCUSSION

Previous studies [24][25] have shown that SVM with Normalized Poly Kernel achieved the best performance compared to other kernels. However, in other studies [9][17], SVM with ordinary kernels, linear and RBF kernel, showed less effective performance than other algorithms. It is because a key success factor in SVM is the choice of kernel function. To master the learning ability, SVM requires a suitable kernel function. Besides, The use of appropriate kernel functions can drastically lower computational efforts [24]. Therefore, in the current study, we employed SVM with Normalized Poly kernel.

The results have shown that SVM with Normalized Poly Kernel outperformed MNB both on data of the economic aspects and the general aspects categorized into two-classes and three-classes dataset. SVM obtained the highest performance on general aspect data using two classes dataset in terms of average accuracy, precision, recall, and f-measure with the value of 82.00%, 82.24%, 82.01%, and 81.84%, respectively. SVM with Normalized Poly Kernel was proven to have an excellent performance in doing a sentiment analysis. Unfortunately, in the current, we did not compare each kernel.

Both algorithms, SVM and MNB, achieved better performance using two-classes than three-classes dataset (see Table 2 to Table 4). This means that the algorithms still find it challenging to classify neutral sentiments. Further research is needed to find a suitable method that can overcome the problem.

The built algorithms were evaluated and validated using confusion matrix and ten-fold cross-validation. Ten-fold cross-validation was employed to validate the robustness of the algorithms. The value of the standard deviation obtained in Table 2, Table 3, Table 4, and Table 5 was less than 1, which means that the performance results of the algorithms are relatively low in variation. Therefore, it could be concluded that the built model can generalize new data accurately, and it can avoid overfitting as well.

Regarding Fig. 6a and 6b, most Indonesians agree with the government's policies in handling COVID-19 in terms of economic aspects. As for general aspects, most Indonesians were not satisfied with the government's performance in dealing with COVID-19. On the economic aspect, the words that most often appeared are *dampak ekonomi*, *dampak covid*, *virus corona*, *pemerintah*, *Indonesia*, *sosial ekonomi*, and *dampak sosial*". While the words that often appeared in general aspects are *pemerintah*, *lockdown*, *Indonesia*, *rakyat*, *masyarakat*, *covid*, and *covid Indonesia*. Based on the experimental results, this research work can be used as a decision support system for the government to improve its performance in dealing with COVID-19 by looking at the public opinions.

The limitation of this study is that there many irrelevant features involved in the classification process so that the algorithms need a big effort to process the data.

VI. CONCLUSIONS

The results have shown that most Indonesians agreed with government policies in dealing with the COVID-19's economic impacts. As for general aspects, most Indonesians were not satisfied with the government's performance in handling COVID-19. In this study, we have also built machine learning algorithms to predict the sentiment analysis of the unseen data. We have compared two algorithms, SVM with Normalized Poly Kernel and MNB. The results have shown that the SVM algorithm with the Normalized Poly Kernel is the best algorithm in predicting sentiments and outperformed MNB in all test models. The SVM provided the highest accuracy in sentiments based on general aspects using two classes with an average accuracy of 82.00%, the precision of 82.24%, the recall of 82.01%, and the f-measure of 81.84%. Therefore, SVM can be used as an intelligent algorithm to conduct a sentiment analysis for new data. Additionally, SVM and MNB are very robust in analyzing data with two classes (positive and negative). The results of this sentiment analysis also can be used as a decision support system for the Indonesian government to improve its performance in handling COVID-19. For future works, researchers can combine the built model with feature selection to reduce the dimensionality, remove irrelevant features, select valuable features, and reduce computational time to improve the classification models' performance.

REFERENCES

- [1] 2019. Katadata, "Berapa media sosial indonesia?," katadata.id. pengguna https://databoks.katadata.co.id/datapublish/2019/02/08/berapa-pengguna-media-sosial-indonesia (accessed May 05, 2020). pesat [2] "Pengguna Twitter indonesia 2018." 2019. AntaraNews, tumbuh di AntaraNews.com.
- https://www.antaranews.com/berita/839825/pengguna-twitter-indonesia-tumbuh-pesat-pada-2018 (accessed May 05, 2020).
- [3] Gugus Percepatan Penanganan COVID-19, "Data Covid-19 di Indonesia," *covid19.go.id*, 2020. https://covid19.go.id/ (accessed May 14, 2020).
- [4] Suara, "Pemerintah Indonesia dinilai lambat mengantisipasi covid-19 sejak dini," suara.com, 2020. https://www.suara.com/news/2020/04/10/025500/pemerintah-indonesia-dinilai-lambat-mengantisipasi-covid-19-sejak-dini?page=all (accessed May 14, 2020).
- [5] DetikNews, "Pemerintah dinilai lambat tangani corona, Jokowi: Kita tak ingin grasa-grusu," news.detik.com, 2020. https://news.detik.com/berita/d-4987368/pemerintah-dinilai-lambat-tangani-corona-jokowi-kita-tak-ingin-grasa-grusu (accessed May 14, 2020).
- [6] I. Zulfa and E. Winarko, "Sentimen analisis tweet berbahasa Indonesia dengan deep belief network," IJCCS (Indonesian J. Comput. Cybern. Syst., vol. 11, no. 2, pp. 187, 2017, doi: 10.22146/ijccs.24716.
- [7] J. Michie, "The covid-19 crisis and the future of the economy and economics," Int. Rev. Appl. Econ., vol. 34, no. 3, pp. 301–303, 2020, doi: 10.1080/02692171.2020.1756040.
- [8] R. Shahid, S. T. Javed, and K. Zafar, "Feature selection based classification of sentiment analysis using biogeography optimization algorithm," in 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), 2017, pp. 1–5, doi: 10.1109/ICIEECT.2017.7916549.
- [9] D. A. Kristiyanti, A. H. Umam, M. Wahyudi, R. Amin, and L. Marlinda, "Comparison of SVM Naïve Bayes Algorithm for sentiment analysis toward West Java Governor candidate period 2018-2023 based on public opinion on Twitter," in 2018 6th International Conference on Cyber and IT Service Management, CITSM 2018, 2018, pp. 1–6, doi: 10.1109/CITSM.2018.8674352.
- [10] G. Singh, B. Kumar, L. Gaur, and A. Tyagi, "Comparison between Multinomial and Bernoulli Naïve Bayes for Text Classification," in 2019 International Conference on Automation, Computational and Technology Management, ICACTM, 2019, pp. 593–596, doi: 10.1109/ICACTM.2019.8776800.

- [11] A. A. Lutfi, A. E. Permanasari, and S. Fauziati, "Sentiment analysis in the sales review of Indonesian marketplace by utilizing Support Vector Machine," J. Inf. Syst. Eng. Bus. Intell., vol. 4, no. 2, p. 169, 2018, doi: 10.20473/jisebi.4.2.169.
- [12] A. M. Rahat, A. Kahir, and A. K. M. Masum, "Comparison of Naive Bayes and SVM Algorithm based on sentiment analysis using review dataset," in 8th International Conference on System Modeling & Advancement in Research Trends, 2020, pp. 266-270, doi: 10.1109/smart46866.2019.9117512.
- S. Rana and A. Singh, "Comparative analysis of sentiment orientation using SVM and Naive Bayes techniques," in 2nd International [13] Conference on Next Generation Computing Technologies, 2016, pp. 106–111, doi: 10.1109/NGCT.2016.7877399. U. Kumari, A. K. Sharma, and D. Soni, "Sentiment analysis of smart phone product review using SVM classification technique," in
- [14] International Conference on Energy, Communication, Data Analytics and Soft Computing, 2017, pp. 1469-1474.
- [15] B. Shamantha Rai, S. M. Shetty, and P. Rai, "Sentiment analysis using machine learning classifiers: evaluation of performance," in 2019 IEEE 4th International Conference on Computer and Communication Systems, ICCCS, 2019, pp. 21-25, doi: 10.1109/CCOMS.2019.8821650.
- M. Wongkar and A. Angdresey, "Sentiment analysis using Naive Bayes Algorithm of the data crawler: Twitter," in Proceedings of [16] 2019 4th International Conference on Informatics and Computing, ICIC, 2019, pp. 1-5, doi: 10.1109/ICIC47613.2019.8985884.
- [17] P. Juneja and U. Ojha, "Casting online votes: To predict offline results using sentiment analysis by machine learning classifiers," in 8th International Conference on Computing, Communications and Networking Technologies, ICCCNT 2017, 2017, pp. 1-6, doi: 10.1109/ICCCNT.2017.8203996.
- M. I. Zul, F. Yulia, and D. Nurmalasari, "Social media sentiment analysis using K-means and naïve bayes algorithm," in Proceedings -[18] 2018 2nd International Conference on Electrical Engineering and Informatics, 2018, pp. 24–29, doi: 10.1109/ICon-EEI.2018.8784326.
- [19] S. L. Ramdhani, R. Andreswari, and M. A. Hasibuan, "Sentiment analysis of product reviews using Naive Bayes Algorithm: A Case Study," in Proceedings - 2nd East Indonesia Conference on Computer and Information Technology: Internet of Things for Industry, EIConCIT, 2018, pp. 123-127, doi: 10.1109/EIConCIT.2018.8878528.
- L. Muflikhah, D. J. Haryanto, A. A. Soebroto, and E. Santoso, "High performance of polynomial kernel at SVM Algorithm for sentiment analysis," J. Inf. Technol. Comput. Sci., vol. 3, no. 2, p. 194, 2018, doi: 10.25126/jitecs.20183260. [20]
- [21] M. M. Hossain and M. S. Miah, "Evaluation of different SVM kernels for predicting customer churn," in 2015 18th International Conference on Computer and Information Technology, ICCIT 2015, 2016, pp. 1-4, doi: 10.1109/ICCITechn.2015.7488032.
- [22] S. Arafin Mahtab, N. Islam, and M. Mahfuzur Rahaman, "Sentiment Analysis on Bangladesh Cricket with Support Vector Machine," in 2018 International Conference on Bangla Speech and Language Processing, ICBSLP, 2018, pp. 21-22, doi: 10.1109/ICBSLP.2018.8554585.
- [23] J. A. Banados and K. J. Espinosa, "Optimizing Support Vector Machine in classifying sentiments on product brands from Twitter," in IISA 2014 - 5th International Conference on Information, Intelligence, Systems and Applications, 2014, pp. 75-80, doi: 10.1109/IISA.2014.6878768.
- S. K. Trivedi and S. Dey, "Effect of Various Kernels and Feature Selection Methods on SVM Performance for Detecting Email [24] Spams," Int. J. Comput. Appl., vol. 66, no. 21, pp. 18-23, 2013, doi: 10.5120/11240-6433.
- A. Alalshekmubarak, A. Hussain, and Q.-F. Wang, "Off-Line Handwritten Arabic Word Recognition Using SVMs with Normalized [25] Poly Kernel," in Neural Information Processing, 2012, pp. 85-91.
- [26] M. O. Pratama et al., "The sentiment analysis of Indonesia commuter line using machine learning based on twitter data," J. Phys. Conf. Ser., vol. 1193, no. 1, pp. 1-6, 2019, doi: 10.1088/1742-6596/1193/1/012029.
- Taspinar, "Twitterscraper Library." https://github.com/taspinar/twitterscraper (accessed May 07, 2020). [27]
- [28] D. H. Wahid and A. SN, "Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity," Indonesian J. Comput. Cybern. Syst., 10, 207–218, 2016, vol. no. 2. [Online]. Available: pp. https://journal.ugm.ac.id/ijccs/article/view/16625/11694.
- Sastrawi, "Sastrawi Library." https://pypi.org/project/Sastrawi/ (accessed May 05, 2020). [29]
- [30] P. P. M. Surya, L. V. Seetha, and B. Subbulakshmi, "Analysis of user emotions and opinion using Multinomial Naive Bayes Classifier," in Proceedings of the 3rd International Conference on Electronics and Communication and Aerospace Technology, ICECA 2019, 2019, pp. 410-415, doi: 10.1109/ICECA.2019.8822096.
- [31] T. Tokunaga, T. Tokunaga, I. Makoto, and I. Makoto, "Text categorization based on weighted inverse document frequency," Spec. Inf. Process (SIG-IPSJ), 1994, Interes. Groups Soc. Japan [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.49.7015.
- J. Ren, S. D. Lee, X. Chen, B. Kao, R. Cheng, and D. Cheung, "Naive bayes classification of uncertain data," in Proceedings IEEE [32] International Conference on Data Mining, ICDM, 2009, no. September 2014, pp. 944–949, doi: 10.1109/ICDM.2009.90.
- [33] E. Prasetyo, Data Mining - mengolah data menjadi informasi menggunakan Matlab. Penerbit Andi, 2014.
- [34] I. H. Witten, E. Frank, and M. A. Hall, Data Mining Practical Machine Learning Tools and Techniques, Third Edit. Burlington: Elsevier, 2010.